Amendments to the Claims

Please amend Claims 1, 4-7 and 9-13. Please add new Claims 14 and 15. The Claim Listing below will replace all prior versions of the claims in the application:

Claim Listing

1. (Currently amended) A method of identifying one or more proteins in an unannotated DNA sequence, the method comprising:
   (a) dividing the DNA sequence into a plurality of sequence fragments each fragment being of substantially the same length and from about 300 to 5000 base pairs long;
   (b) performing a six frame translation of each of the DNA sequence fragments to obtain six translated amino acid sequence fragments for each DNA sequence fragment;
   (c) subjecting each of the translated sequence fragments to theoretical digestion to obtain a plurality of cleaved peptide sequences; and
   (d) comparing experimental empirical data for peptide fragments from a protein digested in the same manner as the theoretical digestion at step (c) with the theoretical data generated in step (c) for each of the translated sequence fragments to identify one or more translated sequence fragments which include a significant number of peptides present in the digested protein.

2. (Original) The method of claim 1 wherein the step (a) of dividing the DNA sequence into a plurality of sequence fragments is performed before the step (b) of performing the six frame translation.

3. (Original) The method of claim 1 wherein the step (a) of dividing the DNA sequence into a plurality of sequence fragments is performed after the step (b) of performing the six frame translation.

4. (Currently amended) The method of ~~any preceding~~ claim 1 wherein theoretically generated peptide masses are compared to the masses of the peptides experimentally

generated by the digested protein and the sequence fragment which has the greatest number of theoretical peptide masses correlating to the empirical data ~~indicates~~ is identified as indicating the likely location of the protein of interest in the DNA sequence.

5. (Currently amended) The method of ~~any preceding~~ claim 1 wherein the masses of the peptides experimentally generated from the digested protein are determined by mass spectrometry.

6. (Currently amended) The method of ~~any preceding~~ claim 1 wherein the DNA sequence is duplicated into a duplicate and an original and the original and duplicate are split in such a manner that the sequence fragments from the ~~original~~ duplicate overlap divisions in the original genome sequence.

7. (Currently amended) The method of ~~any preceding~~ claim 1 wherein the sequence fragments are from 800 to 1200 base pairs long.

8. (Original) The method of claim 7 wherein the sequence fragments are around 1000 to 1050 bases long.

9. (Currently amended) The method of ~~any preceding~~ claim 1 wherein steps (c) and (a) are performed twice using different enzymes and ~~the~~ data ~~is~~ from the two digests is combined and analysed to identify the protein coding region of interest.

10. (Currently amended) The method of ~~any preceding~~ claim 1 wherein the in theoretical digest of step (c) all theoretical peptides which contain a stop codon are discarded.

11. (Currently amended) The method of ~~any preceding~~ claim 1 wherein the fragments are numbered so that an overlapping fragment is numbered n where the fragments it overlaps are numbered n-1 and n ~~plus~~ $\pm$ 1, where n is an integer.

12. (Currently amended) A method of identifying one or more proteins in unannotated DNA sequence, the method comprising:

(a) performing a six frame translation of a DNA sequence to provide six translated amino acid sequences;

(b) dividing the six translated amino acid sequences into a plurality of fragments, each fragment comprising 100-1666 amino acids;

(c) subjecting each of the fragments to theoretical digestion to obtain a plurality of cleaved peptide sequences; and

(d) comparing experimental empirical data for peptide fragment for peptide fragments from a protein digested in the same manner as the theoretical digestion at step (c) with theoretical data generated in step (c) for each of the fragments to identify one or more fragments which include a significant number of peptides present in the empirically digested protein.

13. (Currently amended) The method of claim 12 wherein each six translated amino acid sequences is duplicated into an original and a duplicate copy and the original and duplicate of each are split in such a manner that the sequence fragments from the original overlap divisions in the original sequence.

14. (New) The method of claim 12 wherein theoretically generated peptide masses are compared to the masses of the peptides experimentally generated by the digested protein and the sequence fragment which has the greatest number of theoretical peptide masses correlating to the empirical data is identified as indicating the likely location of the protein of interest in the DNA sequence.

15. (New) The method of claim 12 wherein step (c) is performed twice using different enzymes and data from the two digests is combined and analysed to identify a protein coding region of interest.